

Amendments to the Claims:

This listing of claims will replace all prior versions, and listings, of claims in the application:

Listing of Claims:

1.-3. (Cancelled)

4. (Currently Amended) The load balancing method according to claim 9~~claim 1~~, wherein the steps (a) through (d) are sequentially performed when a transaction processing request has been received from a terminal,
wherein, in the step (d), one of the plurality of computers is selected based on the load indexes as an execution computer processing the transaction processing request.

5. (Currently Amended) The load balancing method according to claim 9~~claim 2~~, wherein, in the step (a.1), a number of in-process transactions and a number of job processing processes staying in a CPU system of the computer are measured as the load data measured at constant intervals.

6. (Currently Amended) The load balancing method according to claim 9~~claim 2~~, wherein, in the step (a.1), a number of in-process transactions and a CPU utilization of the computer are measured as the load data measured at constant intervals.

7.-8. (Cancelled)

9. (Currently Amended) A load balancing method in a system comprising a plurality of computers for processing transaction processing requests originating from a plurality of terminals, comprising steps of:

a) estimating load states of respective ones of the plurality of computers;

b) determining estimated elongation rates of processing time for respective ones of the plurality of computers based on the estimated load states, the estimated elongation rates being estimates of a ratio of job processing response time, including queuing time, to total processing time;

c) calculating load indexes of respective ones of the plurality of computers based on the estimated elongation rates; and

d) determining load distribution among the plurality of computers based on the load indexes;

wherein the step (a) comprises the steps of:

a.1) measuring load data of each of the plurality of computers at constant intervals;
and

a.2) estimating a load state of each of the plurality of computers based on the load data;

wherein,

in the step (a), a current number of in-process transactions in each of the plurality of computers is measured in response to start and termination of transaction processing at each of the plurality of computers,

in the step (b), an estimated elongation rate of processing time for each of the plurality of computers is determined based on the estimated current load state and the current number of in-process transactions in the computer; and

The load balancing method according to claim 3, wherein the step (b) comprises the steps of:

———b.1) correcting the estimated load states using current numbers of in-process transactions for respective ones of the plurality of computers to produce corrected estimated load states; and

b.2) determining said estimated elongation rates of processing time for respective ones of the plurality of computers based on the corrected estimated current load states.

10. (Cancelled)

11. (Currently Amended) The load balancing method according to claim 9~~claim 1~~, wherein, in the step (c), an estimated elongation rate is used as a load index of each of the plurality of computers.

12. (Currently Amended) The load balancing method according to claim 9~~claim 1~~, wherein, in the step (c), an estimated elongation rate is one of a before-scheduling estimated elongation rate and an after-scheduling estimated elongation rate,

wherein the before-scheduling estimated elongation rate is an estimated elongation rate calculated from a corresponding estimated load state before the transaction processing request is allocated to respective ones of the plurality of computers, and

the after-scheduling estimated elongation rate is an estimated elongation rate calculated from a corresponding estimated load state after the transaction processing request is allocated to respective ones of the plurality of computers.

13. (Previously Presented) The load balancing method according to claim 12, wherein the step (c) comprises the steps of:

c.1) multiplying the estimated elongation rate of each of the plurality of computers by a current number of in-process transactions in the computer to produce a total estimated elongation rate of the computer; and

c.2) determining the total estimated elongation rate as a load index of the computer.

14. (Previously Presented) The load balancing method according to claim 12, wherein the step (c) comprises the steps of:

c.1) multiplying the estimated elongation rate of each of the plurality of computers by a current number of in-process transactions in the computer to produce a total estimated elongation rate of the computer;

c.2) calculating a total estimated elongation rate difference between an after-scheduling total estimated elongation rate and a before-scheduling total estimated elongation rate of each of the plurality of computers; and

c.3) determining the total estimated elongation rate difference as a load index of the computer.

- 15.-17. (Cancelled)

18. (Currently Amended) The load balancing system according to claim 23~~claim 15~~, wherein the execution computer selector is started up when a transaction processing request has been received from a terminal and then selects one of the plurality of

computers based on the load indexes as an execution computer processing the transaction processing request.

19. (Currently Amended) The load balancing system according to claim 23~~claim 16~~, wherein the load estimator measures a number of in-process transactions and a number of job processing processes staying in a CPU system of the computer as the load data measured at constant intervals.
20. (Currently Amended) The load balancing system according to claim 23~~claim 16~~, wherein the load estimator measures a number of in-process transactions and a CPU utilization of the computer as the load data measured at constant intervals.
- 21.-22. (Cancelled)
23. (Currently Amended) A load balancing system comprising:
a plurality of terminals, each of said plurality of terminals being adapted to originate a transaction processing request;
a plurality of computers, each of said plurality of computers being adapted to process a plurality of transaction processing requests originating from a plurality of terminals in parallel;
a load estimator for estimating load states of respective ones of the plurality of computers;
a load data memory for storing the estimated load states; and
an execution computer selector for selecting one of the plurality of computers as an execution computer to be put in charge of processing a transaction processing request based on load indexes of respective ones of the plurality of computers, wherein the load indexes are calculated from estimated elongation rates of processing time for respective ones of the plurality of computers, wherein the estimated elongation rates are determined based on the estimated load states and are estimates of a ratio of job processing response time, including queuing time, to total processing time;

wherein the load estimator measures load data of each of the plurality of computers at constant intervals and estimates a load state of each of the plurality of computers based on the load data;

wherein the load estimator measures a current number of in-process transactions in each of the plurality of computers in response to start and termination of transaction processing at each of the plurality of computers;

wherein the execution computer selector determines an estimated elongation rate of processing time for each of the plurality of computers based on the estimated current load state and the current number of in-process transactions in the computer; and

~~The load balancing system according to claim 17;~~

wherein the execution computer selector corrects the estimated load states using the current numbers of in-process transactions for respective ones of the plurality of computers to produce corrected estimated load states, and determines estimated elongation rates of processing time for respective ones of the plurality of computers based on the corrected estimated current load states.

24. (Currently Amended) A load balancing system comprising:

a plurality of terminals, each of said plurality of terminals being adapted to originate a transaction processing request;

a plurality of computers, each of said plurality of computers being adapted to process a plurality of transaction processing requests originating from a plurality of terminals in parallel;

a load estimator for estimating load states of respective ones of the plurality of computers;

a load data memory for storing the estimated load states; and

an execution computer selector for selecting one of the plurality of computers as an execution computer to be put in charge of processing a transaction processing request based on load indexes of respective ones of the plurality of computers, wherein the load indexes are calculated from estimated elongation rates of processing time for respective ones of the plurality of computers, wherein the estimated elongation rates

are determined based on the estimated load states and are estimates of a ratio of job processing response time, including queuing time, to total processing time;

wherein the load estimator measures load data of each of the plurality of computers at constant intervals and estimates a load state of each of the plurality of computers based on the load data;

wherein the load estimator measures a current number of in-process transactions in each of the plurality of computers in response to start and termination of transaction processing at each of the plurality of computers;

wherein the execution computer selector determines an estimated elongation rate of processing time for each of the plurality of computers based on the estimated current load state and the current number of in-process transactions in the computer;

wherein the load estimator measures load data of each of the plurality of computers at constant intervals to produce a sequence of load data, and estimates a current load state of each of the plurality of computers based on the sequence of load data; and

The load balancing system according to claim 22,

wherein the execution computer selector corrects the estimated load states using the current numbers of in-process transactions for respective ones of the plurality of computers to produce corrected estimated load states, and determines estimated elongation rates of processing time for respective ones of the plurality of computers based on the corrected estimated current load states.

25. (Currently Amended) The load balancing system according to claim 23~~claim 15~~, wherein the execution computer selector uses an estimated elongation rate as a load index of each of the plurality of computers.

26. (Currently Amended) The load balancing system according to claim 23~~claim 15~~, wherein an estimated elongation rate is one of a before-scheduling estimated elongation rate and an after-scheduling estimated elongation rate,

wherein the before-scheduling estimated elongation rate is an estimated elongation rate calculated from a corresponding estimated load state before the

transaction processing request is allocated to respective ones of the plurality of computers, and

the after-scheduling estimated elongation rate is an estimated elongation rate calculated from a corresponding estimated load state after the transaction processing request is allocated to respective ones of the plurality of computers.

27. (Previously Presented) The load balancing system according to claim 26, wherein the execution computer selector multiplies the estimated elongation rate of each of the plurality of computers by a current number of in-process transactions in the computer to produce a total estimated elongation rate of the computer, and determines the total estimated elongation rate as a load index of the computer.
28. (Previously Presented) The load balancing system according to claim 26, wherein the execution computer selector multiplies the estimated elongation rate of each of the plurality of computers by a current number of in-process transactions in the computer to produce a total estimated elongation rate of the computer, calculates a total estimated elongation rate difference between an after-scheduling total estimated elongation rate and a before-scheduling total estimated elongation rate of each of the plurality of computers, and determines the total estimated elongation rate difference as a load index of the computer.
29. (Currently Amended) The load balancing system according to claim 23~~claim 15~~, further comprising:
 - a control node connected between the terminals and the plurality of computers, wherein the control node includes the load estimator, the load data memory, and the execution computer selector.
30. (Currently Amended) The load balancing system according to claim 23~~claim 15~~, further comprising:
 - a switching device connecting the plurality of computers to each other,
 - wherein each of the plurality of computers includes the load estimator, the load data memory, and the execution computer selector, wherein an estimated load state

estimated at a computer is transferred to all other computers through the switching device so that the load state memory of each of the plurality of computers stores same load state data.

31. (Currently Amended) A load balancing system comprising:
a plurality of terminals, each of said plurality of terminals being adapted to originate a transaction processing request;
a plurality of computers, each of said plurality of computers being adapted to process a plurality of transaction processing requests originating from a plurality of terminals in parallel;
a load estimator for estimating load states of respective ones of the plurality of computers;
a load data memory for storing the estimated load states;
an execution computer selector for selecting one of the plurality of computers as an execution computer to be put in charge of processing a transaction processing request based on load indexes of respective ones of the plurality of computers, wherein the load indexes are calculated from estimated elongation rates of processing time for respective ones of the plurality of computers, wherein the estimated elongation rates are determined based on the estimated load states and are estimates of a ratio of job processing response time, including queuing time, to total processing time; and
a switching device connecting the plurality of computers to each other,
wherein each of the plurality of computers includes the load estimator, the load data memory, and the execution computer selector, wherein an estimated load state estimated at a computer is transferred to all other computers through the switching device so that the load state memory of each of the plurality of computers stores same load state data; and

~~The load balancing system according to claim 30,~~

wherein the execution computer selector selects a computer to which the execution computer selector belongs when the load index of the computer is smaller than a value obtained by multiplying a minimum load index among the plurality of computers by a predetermined factor greater than 1.

32. (Currently Amended) The load balancing system according to claim 23~~claim 15~~, further comprising:

a switching device connecting the plurality of computers to each other, the switching device including the load data memory which is shared among the plurality of computers,

wherein each of the plurality of computers includes the load estimator and the execution computer selector, wherein an estimated load state estimated at a computer is transferred to the switching device to be stored in the load state memory.

33. (Previously Presented) A load balancing system comprising:

a plurality of terminals, each of which originates said plurality of terminals being adapted to originate a transaction processing request;

a plurality of computers, each of which processes said plurality of computers being adapted to process a plurality of transaction processing requests originating from a plurality of terminals in parallel;

a load estimator for estimating load states of respective ones of the plurality of computers;

a load data memory for storing the estimated load states; and

an execution computer selector for selecting one of the plurality of computers as an execution computer to be put in charge of processing a transaction processing request based on load indexes of respective ones of the plurality of computers, wherein the load indexes are calculated from estimated elongation rates of processing time for respective ones of the plurality of computers, wherein the estimated elongation rates are determined based on the estimated load states;

the system further comprising:

an interim control node connected between the plurality of terminals and the plurality of computers, wherein the interim control node includes an interim execution computer selector for selecting one of the plurality of computers according to a predetermined rule; and

a switching device connecting the plurality of computers to each other,

wherein each of the plurality of computers includes the load estimator, the load data memory, and the execution computer selector, wherein an estimated load state estimated at a computer is transferred to all other computers through the switching device so that the load state memory of each of the plurality of computers stores same load state data, and

wherein the execution computer selector selects a computer to which the execution computer selector belongs when the load index of the computer is smaller than a value obtained by multiplying a minimum load index among the plurality of computers by a predetermined factor greater than 1.

34. (Previously Presented) A load balancing system according to claim 15, comprising:
- a plurality of terminals, each of which originates said plurality of terminals being adapted to originate a transaction processing request;
 - a plurality of computers, each of which processes said plurality of computers being adapted to process a plurality of transaction processing requests originating from a plurality of terminals in parallel;
 - a load estimator for estimating load states of respective ones of the plurality of computers;
 - a load data memory for storing the estimated load states; and
 - an execution computer selector for selecting one of the plurality of computers as an execution computer to be put in charge of processing a transaction processing request based on load indexes of respective ones of the plurality of computers, wherein the load indexes are calculated from estimated elongation rates of processing time for respective ones of the plurality of computers, wherein the estimated elongation rates are determined based on the estimated load states;
- the system further comprising:
- a switching device connecting the plurality of computers to each other, the switching device including the load data memory which is shared among the plurality of computers,

wherein each of the plurality of computers includes the load estimator and the execution computer selector, wherein an estimated load state estimated at a computer is transferred to the switching device to be stored in the load state memory, and

wherein the execution computer selector selects a computer to which the execution computer selector belongs when the load index of the computer is smaller than a value obtained by multiplying a minimum load index among the plurality of computers by a predetermined factor greater than 1.

35. (Previously Presented) A load balancing method in a system comprising a plurality of computers for processing transaction processing requests originating from a plurality of terminals, comprising steps of:

- a) estimating load states of respective ones of the plurality of computers;
- b) determining estimated elongation rates of processing time for respective ones of the plurality of computers based on the estimated load states;
- c) calculating load indexes of respective ones of the plurality of computers based on the estimated elongation rates; and
- d) determining load distribution among the plurality of computers based on the load indexes,

wherein the estimated elongation rate E is a ratio of a processing time required for a job processing process to a net processing time, wherein the estimated elongation rate E is obtained by the following equation:

$$E = X/(X - P \cdot P), \text{ when } P < N,$$
$$= N + 1.0, \text{ when } P \geq N,$$

where $X = N \cdot (P + 1)$, N is the number of in-process transactions in a computer and P is the number of transaction processes in a CPU system of the computer.

36. (Previously Presented) A load balancing method in a system comprising a plurality of computers for processing transaction processing requests originating from a plurality of terminals, comprising steps of:

- a) estimating load states of respective ones of the plurality of computers;
- b) determining estimated elongation rates of processing time for respective ones of the plurality of computers based on the estimated load states;

c) calculating load indexes of respective ones of the plurality of computers based on the estimated elongation rates; and

d) determining load distribution among the plurality of computers based on the load indexes,

wherein the estimated elongation rate E is obtained by the following equation:

$$E = N(1-R)/(N(1-R)-R \cdot R)$$

where N is a number of in-process transactions in a computer and R is a CPU utilization.

37.-39. (Cancelled)

40. (Previously Presented) A load balancing method in a system comprising an interim load balancing device connecting a plurality of terminals and a plurality of execution computers, comprising steps of:

at the interim load balancing device,

a) setting a predetermined distribution scheme;

b) when receiving a transaction processing request from a terminal, selecting an interim destination execution computer from the plurality of execution computers according to the predetermined distribution scheme;

c) sending the transaction processing request to the interim destination execution computer;

at each of the execution computers,

d) estimating a load state of the execution computer based on load data measured at regular intervals including a number of in-process transactions and one of a CPU utilization and a number of job processing processes staying in a CPU system of the execution computer;

e) storing the load state of the execution computer and a load state received from another execution computer in a load state memory;

f) transferring the load state of the execution computer to all other execution computers;

g) when receiving the transaction processing request from the interim load balancing device, determining estimated elongation rates of processing time for

respective ones of the plurality of execution computers based on the load states stored in the load state memory;

h) calculating load indexes of respective ones of the plurality of execution computers from the estimated elongation rates;

said method further comprising the steps of:

i) determining based on the load indexes whether the transaction processing request should be processed by the interim destination execution computer or transferred to another execution computer;

j) when it is determined that the transaction processing request should be transferred to another execution computer, determining in said interim destination execution computer, a final destination computer among the plurality of execution computers based on the load indexes to transfer it to the final destination computer; and

k) when it is determined in said interim destination execution computer, that the transaction processing request should be processed by the interim destination execution computer, processing the transaction processing request in said interim destination execution computer.

41. (Previously Presented) A load balancing method according to claim 40, wherein, in the step (a), the predetermined distribution scheme is a static distribution method such that the plurality of terminals are previously divided into a plurality of groups and said interim destination execution computer is determined depending on which one of the plurality of groups a request originating terminal belongs to.
42. (Previously Presented) A load balancing method according to claim 40, wherein, in the step (a), the predetermined distribution scheme is a static round distribution method such that the plurality of execution computers are sequentially and repeatedly selected as an interim destination execution computer in the arrival order.
43. (Previously Presented) A load balancing method according to claim 40, wherein the step (a) comprises the steps of:

a.1) receiving load data from each of the plurality of execution computers at regular intervals, the load data including a number of in-process transactions and one of a CPU utilization and a number of job processing processes staying in a CPU system of the execution computer;

a.2) determining the predetermined distribution scheme based on the load data so that a transaction processing load is balanced among the plurality of execution computers.

44.-45. (Cancelled)

46. (Previously Presented) A recording medium storing:
a first computer program for instructing a first computer of an interim load balancing device to distribute transaction processing requests originating from a plurality of terminals to a plurality of execution computers; and
a second computer program for instructing each of the plurality of execution computers to balance a transaction processing load among the plurality of execution computers,

wherein the first computer program comprises the steps of:

a) setting a predetermined distribution scheme;

b) when receiving a transaction processing request from a terminal, determining an interim execution computer according to the predetermined distribution scheme; and

c) sending the transaction processing request to the interim execution computer,

wherein the second computer program comprises the steps of:

d) estimating a load state of the execution computer based on load data measured at regular intervals including a number of in-process transactions and one of a CPU utilization and a number of job processing processes staying in a CPU system of the execution computer;

e) storing the load state of the execution computer and a load state received from another execution computer in a load state memory;

f) transferring the load state of the execution computer to all other execution computers;

g) when receiving the transaction processing request from the interim load balancing device, determining estimated elongation rates of processing time for respective ones of the plurality of execution computers based on the load states stored in the load state memory;

h) calculating load indexes of respective ones of the plurality of execution computers from the estimated elongation rates;

i) determining based on the load indexes whether the transaction processing request should be processed by the execution computer or transferred to another execution computer;

j) when it is determined that the transaction processing request should be transferred to another execution computer, determining a final destination computer among the plurality of execution computers based on the load indexes to transfer it to the final destination computer; and

k) when it is determined that the transaction processing request should be processed by the execution computer, processing the transaction processing request.

47.-49. (Cancelled)